

GUJARAT TECHNOLOGICAL UNIVERSITY**BE - SEMESTER-VII (NEW) EXAMINATION – WINTER 2023****Subject Code:3174207****Date:12-12-2023****Subject Name:Big Data Analytics****Time:10:30 AM TO 01:00 PM****Total Marks:70****Instructions:**

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Simple and non-programmable scientific calculators are allowed.

		MARKS
Q.1	(a) Explain Structured Data, Unstructured Data and Semi- Structured Data.	03
	(b) Distinguish Traditional Vs Big data.	04
	(c) Explain the following terms with suitable example. i) Predictive Analytics ii) Perceptive Analytics	07
Q.2	(a) Define big data. Enlist the characteristics of big data.	03
	(b) Discuss the HDFS architecture in detail.	04
	(c) Explain Map-Reduce framework in detail. Draw the architectural diagram for Physical Organization of Computer Nodes.	07
	OR	
	(c) Explain hadoop environment (i.e. echo system) with major components in detail.	07
Q.3	(a) Explain hadoop scaling out in detail.	03
	(b) Explain NoSQL data architecture patterns family document stores in detail.	04
	(c) Enlist and explain four ways that NoSQL systems handle big data problems.	07
	OR	
Q.3	(a) Write the use and syntax of following HDFS commands: i) chown ii) put iii) get	03
	(b) Define NoSQL database. Explain CAP theorem in detail.	04
	(c) Explain the following NoSQL data architecture patterns families i) Graph store ii) Column store (Bigtable)	07
Q.4	(a) Define data stream. Explain stream computing in detail.	03
	(b) Explain Hive architecture in detail.	04
	(c) Explain Stream Data Model and Architecture in detail.	07
	OR	
Q.4	(a) Explain decaying window in case of mining data stream.	03
	(b) Enlist and explain features of apache Pig	04
	(c) Explain Real Time Analysis Platform (RTAP) in detail with suitable example.	07
Q.5	(a) Distinguish apache Pig Vs. Map-reduce.	03
	(b) Explain different types of tables in Hive.	04
	(c) Explain in-memory computing. Explain how performs in-memory computing for data stream.	07
	OR	
Q.5	(a) Explain partitioning and bucketing in Hive with suitable example.	03
	(b) Explain different data processing operators in apache Pig.	04
	(c) Explain different components/modules of apache Spark.	07
